

CLAIM AMENDMENTS

Claim 1. (currently amended) A method for modifying an audio visual recording originally produced with an original audio track of an original speaker, using a second audio dub track of a second speaker, to produce a new transformed audio visual recording with synchronized audio to facial expressive speech of the second audio dub track spoken by the original speaker, comprising analyzing the original audio track to convert it into ~~phonemes as a time-coded phoneme stream~~ a continuous time coded facial and acoustic symbol stream to identify corresponding visual facial motions of the original speaker to create ~~frames~~ continuous data sets of facial motion corresponding to speech ~~phoneme~~ utterance states and transformations, storing these ~~frames~~ continuous data sets in a database, analyzing the second audio dub track to convert it to ~~phonemes as a time-coded phoneme stream~~ a continuous time coded facial and acoustic symbol stream, using the second audio dub ~~track time-coded phoneme stream~~ track's continuous time coded facial and acoustic symbol stream to animate the original speaker's face, synchronized to the second audio dub track to create natural continuous facial speech expression by the original speaker of the second dub audio track.

Claim 2. (original) The method of Claim 1 wherein said second audio dub track is spoken in a language different from that of the original speaker.

Claim 3. (currently amended) The method of Claim 1 wherein said ~~phonemes~~ continuous time coded facial and acoustic symbols comprise diphones and visemes.

Claim 4. (currently amended) The method of Claim 1 wherein said ~~phonemes~~ continuous time coded facial and acoustic symbols comprise triphones and visemes.

Claim 5. (currently amended) The method of Claim 1 comprising using a set of fixed facial reference points to track the facial transformation from one ~~phoneme~~ continuous time coded facial and acoustic symbol to another ~~phoneme~~ continuous time coded facial and acoustic symbol.

Claim 6. (original) The method of Claim 5 comprising using a computer motion tracking system to record the optical flow path for each fixed reference point.

Claim 7. (currently amended) The method of Claim 1 further comprising using a set of fixed facial reference points to track the facial transformation from one continuous time coded facial and acoustic symbol ~~phoneme~~ to another continuous time coded facial and acoustic symbol ~~phoneme~~.

Claim 8. (currently amended) The method of Claim 7 further comprising accumulating a database of recorded ~~triphone and diphone~~ mouth transformation optical flow paths associated with continuous time coded facial and acoustic symbols.

Claim 9. (original) The method of Claim 5 further comprising adding an emotional elicitation process by using a computer motion tracking system to record a number of facial control points for each emotion.

Claim 10. (currently amended) The method of Claim 9 in which the facial control points comprise, but are not limited to, the chin, ~~outside of the mouth and inside of the lips~~ teeth, inner lips, outer lips, cheeks and jowls.

Claim 11. (currently amended) The method of Claim 7 comprising accumulating a database of visemes by recording the facial control points corresponding to each ~~phoneme~~ continuous time coded facial and acoustic symbol.

Claim 12. (original) The method of Claim 11 comprising accumulating a database of muzzle patches by mapping the facial control points for each viseme.

Claim 13. (currently amended) The method of Claim 12 comprising selecting muzzle patches from the speaker's viseme database based on the second dub audio track and ~~phoneme~~ continuous time coded facial and acoustic symbol to viseme sequence and applying the selected muzzle patches onto a three dimensional facial muzzle model.

Claim 14. (original) The method of Claim 13 further comprising also collecting a set of patches of light, color and texture from the sampled first speaker's muzzle patches.

Claim 15. (original) The method of Claim 1 in which the fixed reference points for both speakers are mapped using radar.

Claim 16. (currently amended) The method of Claim 15 in which the dub audio track radar measurements are referenced to a particular continuous time coded facial and acoustic symbol phoneme or phoneme to phoneme or symbol to symbol transitions in time.

Claim 17. (currently amended) A method for modifying an audio visual recording originally produced with an original audio track of an original screen actor, using a second audio dub track of a second screen actor, to produce a new transformed audio visual recording with synchronized audio to facial expressive speech of the second audio dub track spoken by the original screen actor, comprising analyzing the original audio track to convert it into ~~phonemes as a time-coded phoneme stream~~ a continuous time coded facial and acoustic symbol stream to identify corresponding visual facial motions of the original speaker to create ~~frames~~ continuous data sets of facial motion corresponding to speech ~~phoneme~~ utterance states and transformations, storing these ~~frames~~ continuous data sets in a database, analyzing the second audio dub track to convert it to ~~phonemes as a time-coded phoneme stream~~ a continuous time coded facial and acoustic symbol stream, using the second audio dub ~~track time-coded phoneme stream~~ track's continuous time coded facial and acoustic symbol stream to animate the original screen actor's face, synchronized to the second audio dub track to create natural continuous facial speech expression by the original screen actor of the second dub audio track.

Claim 18. (cancelled)

Claim 19. (currently amended) The method of Claim 17 comprising using a computer vision system to track and record a database of visemes of the actors' head position, facial motions of

the jaw, and the lip motion during speech for each ~~frame~~ continuous time coded facial and acoustic symbol.

Claim 20. (original) The method of Claim 19 comprising tracking fixed reference control points on the head, jaw and lips.

Claim 21. (currently amended) The method of Claim 20 in which the reference control points comprise the outside edge of the lips, the inside edge of the lips, the edge of the chin, the tip of the tongue, the bottom edge of the upper teeth, the upper edge of the lower teeth, the cheeks and jowls, the nose and the eyes.

Claim 22. (original) The method of Claim 21 comprising accumulating a database of muzzle patches by mapping the facial control points for each viseme.

Claim 23. (currently amended) The method of Claim 22 comprising selecting muzzle patches from the speaker's viseme database based on the second dub audio track and ~~phoneme~~ continuous time coded facial and acoustic symbol to viseme sequence and applying the selected muzzle patches onto a three dimensional facial muzzle model.

Claim 24. (original) The method of Claim 23 further comprising also collecting a set of patches of light, color and texture from the sampled first speaker's muzzle patches and applying these patches based on the second dub audio track.

Claim 25. (original) The method of Claim 20 in which the fixed reference points for both actors are mapped using radar.

Claim 26. (currently amended) The method of Claim 25 in which the dub audio track radar measurements are referenced to a particular ~~phoneme or phoneme to phoneme transition in time~~
continuous time coded facial and acoustic symbol or symbol to symbol transition.

Claim 27. (original) The method of Claim 25 in which the radar mapping information is transferred to an animation control mixer.

Claim 28. (original) The method of Claim 25 in which the viseme database information is transferred to an animation control mixer.

Claim 29. (original) The methods of Claim 1 or 17 comprising using head modeling to aid in positioning three dimensional visemes.

Claim 30. (currently amended) The methods of Claim 1 or 17 comprising modeling multiple floating point control vertices relative to a generic head shape and relative to on a variety of
~~different~~ speech viseme standards.

Claim 31. (original) The methods of Claim 1 or 17 comprising texture sampling target footage to texture match visemes.

Claim 32. (currently amended) A method for modifying an audio visual recording originally produced with an original audio track of an original screen actor, using a second audio dub track of a second screen actor, to produce a new transformed audio visual recording with synchronized audio to facial expressive speech of the second audio dub track spoken by the original screen actor, comprising analyzing the original audio track to convert it into ~~phonemes as a time-coded phoneme stream~~ a continuous time coded facial and acoustic symbol stream, identifying corresponding visemes of the original screen actor, using radar to measure a set of facial reference points corresponding to speech ~~phoneme~~ continuous time coded facial and acoustic symbol utterance states and transformations, storing the data obtained in a database, analyzing the second audio dub track to convert it to ~~phonemes~~ a continuous time coded facial and acoustic symbol stream ~~as a time-coded phoneme stream~~, identifying corresponding visemes of the second screen actor, using radar to measure a set of facial reference points corresponding to speech ~~phoneme~~ facial and acoustic symbol utterance states and transformations, storing the data obtained in a database, using the second audio dub track time-coded ~~phoneme~~ continuous time coded facial and acoustic symbol stream and the actors visemes to animate the original screen actor's face, synchronized to the second audio dub track to create natural continuous facial speech expression by the original screen actor of the second dub audio track.

Claim 33. (original) The method of Claim 32 in which the database resides in an animation control mixer.

Claim 34. (currently amended) The method of Claim 32 in which the facial reference points comprise the outside edge of the lips, the inside edge of the lips, the edge of the chin, the tip of

the tongue, the bottom edge of the upper teeth, the upper edge of the lower teeth, the cheeks and jowls, the nose and the eyes.

Claim 35. (original) The method of Claim 32 comprising texture sampling target footage to texture match visemes.